June 1965

A MULTIVARIATE TEST OF THE NULL HYPOTHESIS THAT TWO POPULATIONS DO NOT DIFFER IN THE DIRECTION OF THEIR MEAN VECTORS

Charles H. Kraft, Ingram Olkin, Constance van Eeden

Technical Report No. 57

University of Minnesota

Minneapolis, Minnesota

A multivariate test of the null hypothesis that two populations do not differ in the direction of their mean vectors

by

Charles H. Kraft,**Ingram Olkin, Constance van Eeden**
University of Minnesota, Stanford University, University of Minnesota

0. Summary

In this paper the likelihood ratio tests for two multivariate hypotheses are given. The need for these tests arose from a problem in the evaluation of drugs, where the effect of a drug was measured by the change in a 9-component vector; the application to this problem is discussed. The tests are given in a form so that a reader, familiar with matrix calculation, can compute them.

1. Introduction

Let

$$\xi = (\xi_1, \ldots, \xi_p) \quad \text{and} \quad \eta = (\eta_1, \ldots, \eta_p)$$

be the means of two multivariate populations. In many cases an experimenter wants to know if his data are consistent with the hypothesis $\xi = \eta$. When normality is assumed the test of this hypothesis is Hotelling's T^2 (cf e.g. Anderson [1], p. 109) which has power when $\xi \neq \eta$. However there are cases when a test is desired which will have power when ξ and η differ in direction, but does not have power when ξ and η differ only in magnitude. Tests for such situations are derived here.

The particular experiment* from which this problem arose was one conducted to compare two tranquilizers. The experimenter was interested to know whether the two drugs each acted on a different set of symptoms and was not particularly interested to know whether one drug had a greater effect than the other in case

* The authors thank Professor Gordon T. Heistad of the University of Minnesota for bringing this problem to their attention.

** This research was supported in part by the National Science Foundation under Grant Number GP-3816.

the two drugs were affecting the same symptoms. In this case the means ξ and η were the differences between two, pre-treatment and post-treatment, administrations of the Wittenborn psychiatric rating scale [2].

We assume, for the derivation of the tests, that the observations have multivariate normal distributions, say $N(\xi, \Sigma)$ for population 1 and $N(\eta, \Sigma)$ population 2. The null hypothesis H_0 is

$$H_0: \xi = a\eta \quad \text{for some (unknown) number } a.$$

The likelihood ratio test for H_0 against the alternative of arbitrary means will be derived. Also a test is derived for a related hypothesis which may be of interest, namely for

$$H_0': \xi = a\eta + be \quad \text{for some pair of (unknown) numbers } a \text{ and } b,$$

where $e = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$.

2. Definitions

We follow the notation of Anderson [1], and repeat some definitions. Let x_α, $\alpha = 1, \ldots, N_1$, be the (vector) observation made for the α^{th} individual in a sample, sample 1, from population 1, and y_β, $\beta = 1, \ldots, N_2$, that for the β^{th} individual in a sample, sample 2, from population 2. Let $\bar{x}$ be the mean for sample 1 and A_1 the sample cross product matrix. That is

$$(2.1) \qquad \bar{x} = \frac{1}{N_1} \sum_{\alpha=1}^{N_1} x_\alpha = \begin{pmatrix} \frac{1}{N_1} \sum_{\alpha=1}^{N_1} x_{1\alpha} \\ \vdots \\ \frac{1}{N_1} \sum_{\alpha=1}^{N_1} x_{p\alpha} \end{pmatrix} = \begin{pmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_p \end{pmatrix}$$

and A_1 is the matrix whose element in the i^{th} row and j^{th} column is

$$(2.2) \qquad \sum_{\alpha=1}^{N_1} (x_{i\alpha} - \bar{x}_i)(x_{j\alpha} - \bar{x}_j)$$

or

$$(2.3) \qquad A_1 = \sum_{\alpha=1}^{N_1} (x_\alpha - \bar{x})(x_\alpha - \bar{x})' .$$

$\bar{y}$ and A_2 are similarly defined for sample 2.

It is convenient to use the canonical form of the densities of the two samples. To this end let

$$(2.4) \qquad \begin{cases} x = \sqrt{N_1}\,\bar{x} & y = \sqrt{N_2}\,\bar{y} \\ A = A_1 + A_2 & N = N_1 + N_2 \end{cases}$$

Then, since X, Y and A are sufficient statistics for ξ, η and Σ we can work with their density. Let

$$(2.5) \qquad \begin{cases} \mu = \sqrt{N_1}\,\xi & \nu = \sqrt{N_2}\,\eta \\ \quad \Lambda = \Sigma^{-1} , & \end{cases}$$

then the density of X, Y, A is

$$(2.6) \quad p(x,y,A|\mu,\nu,\Lambda) = K|\Lambda|^{\frac{N}{2}} e^{-\frac{1}{2}(x-\mu)\Lambda(x-\mu)' - \frac{1}{2}(y-\nu)\Lambda(y-\nu)' - \frac{1}{2}\operatorname{tr}\Lambda A} .$$

The value, $\hat{\Lambda}$, of Λ which maximizes (2.6) is

$$(2.7) \qquad \hat{\Lambda} = N[A + (x-\mu)'(x-\mu) + (y-\nu)'(y-\nu)]^{-1}$$

and the value of (2.6) for $\Lambda = \hat{\Lambda}$ is

$$(2.8) \quad p(x,y,A|\mu,\nu,\hat{\Lambda}) = Ke^{-\frac{1}{2}pN} N^{\frac{1}{2}pN}|A|^{-\frac{1}{2}N} \left\{1 + (x-\mu)A^{-1}(x-\mu)' + (y-\nu)A^{-1}(y-\nu)'\right\}^{-\frac{1}{2}N}$$

Further (2.8) can be maximized by minimizing

$$(2.9) \qquad (x-\mu)\,A^{-1}(x-\mu)' + (y-\nu)\,A^{-1}(y-\nu)' .$$

If there are no restrictions on (μ,ν) (or equivalently on (ξ,η)) (2.9) is minimized by $\mu = x$ and $\nu = y$, then $\hat{\Lambda} = NA^{-1}$. These are the usual unrestricted maximum likelihood estimates of μ, ν and Λ.

3. The test for H_0

If it is assumed that $\mu = c\nu$, where $c = a\sqrt{\frac{N_1}{N_2}}$, (2.9) becomes

$$(3.1) \qquad (x-c\nu)\, A^{-1}(x-c\nu)' + (y-\nu)\, A^{-1}(y-\nu)'$$

which has a minimum at

$$(3.2) \qquad \begin{cases} \hat{\nu} = \dfrac{\hat{c}x + y}{\hat{c}^2 + 1} \\[2ex] \hat{c} = \dfrac{-yA^{-1}y' + xA^{-1}x' + \sqrt{(yA^{-1}y' - xA^{-1}x')^2 + 4(yA^{-1}x')^2}}{2yA^{-1}x'} \end{cases} .$$

The formula for $\hat{c}$ can be computed from the observations (cf. (2;3) and (2:4)) and, working back, so can that for $\hat{\nu}$, so that the maximum of the density when H_0 is true is

$$(3.3) \qquad p(x,y,A|\hat{c}\hat{\nu}, \hat{\nu}, \hat{\Lambda}) = Ke^{-\frac{1}{2}pN} N^{\frac{1}{2}pN} |A + (x-\hat{c}\hat{\nu})'(x-\hat{c}\hat{\nu}) + (y-\hat{\nu})'(y-\hat{\nu})|^{-\frac{1}{2}N}$$

If there are no restrictions on (μ,ν) the maximum of the density is

$$(3.4) \qquad p(x,y,A|\hat{\mu},\hat{\nu},\hat{\Lambda}) = Ke^{-\frac{1}{2}pN} N^{\frac{1}{2}pN} |A|^{-\frac{1}{2}N}$$

The likelihood ratio test for H_0 is then (cf. Anderson [1], p. 208)

$$(3.5) \qquad T = -2\, \frac{N - 2 - \frac{p}{2} - \frac{1}{2p}}{N} \log \frac{p(x,y,A|\hat{c}\hat{\nu},\hat{\nu},\hat{\Lambda})}{p(x,y,A|\hat{\mu},\hat{\nu},\hat{\Lambda})} =$$

$$= (N - 2 - \frac{p}{2} - \frac{1}{2p}) \log \left\{ 1 + (x-\hat{c}\hat{\nu})\, A^{-1}(x-\hat{c}\hat{\nu})' + (y-\hat{\nu})\, A^{-1}(y-\hat{\nu})' \right\} .$$

T has asymptotically a χ^2-distribution with $p-1$ degrees of freedom. (For an improved approximation see Anderson [1], p. 208 where, for this case, $q_1 = 1 - \frac{1}{p}$ and $q_2 = 1 + \frac{1}{p}$) .

4. The test for H_0'

If it is assumed that $\mu = c\nu + de$, where $c = a\sqrt{\frac{N_1}{N_2}}$ and $d = b\sqrt{N_1}$, (2.9) becomes

$$(4.1) \qquad (x - c\nu - de)\, A^{-1}(x - c\nu - de)' + (y-\nu)\, A^{-1}(y-\nu)'$$

which has a minimum at

$$(4.2) \qquad \begin{cases} \hat{\nu} = \dfrac{y + \hat{c}x - \hat{c}\hat{d}e}{1 + \hat{c}^2} \\[2ex] \hat{d} = \dfrac{e\, A^{-1}x' - \hat{c}e\, A^{-1}y'}{e\, A^{-1}e'} \\[2ex] \hat{c} = \dfrac{xUx' - yUy' + \sqrt{(yUy' - xUx')^2 + 4(xUy')^2}}{2xUy'} \end{cases}$$

where

$$(4.3) \qquad U = A^{-1} - \frac{A^{-1}e'e\, A^{-1}}{e\, A^{-1}e'} \quad .$$

These can be substituted in

$$(4.4) \qquad T' = (N-2-\tfrac{p}{2}-\tfrac{1}{p}) \log \left\{1 + (x - \hat{c}\hat{\nu} - \hat{d}e)\, A^{-1}(x - \hat{c}\hat{\nu} - \hat{d}e)' + (y-\hat{\nu})\, A^{-1}(y-\hat{\nu})'\right\},$$

which has, asymptotically, a χ^2-distribution with p-2 degrees of freedom. (For an improved approximation, see Anderson [1] p. 208 where, for this case, $q_1 = 1 - \frac{2}{p}$ and $q_2 = 1 + \frac{2}{p}$) .

References

[1] Anderson, T. W., An introduction to multivariate statistical analysis, John Wiley, Inc., New York, 1958.

[2] Wittenborn, J. R., Manual: Wittenborn psychiatric rating scales, New York, The Psychological Corporation, 1955.